



## A Short Tour OCR, ICR, DIA

Abdel Belaïd

### ► To cite this version:

Abdel Belaïd. A Short Tour OCR, ICR, DIA. International Workshop on Technology Development in Indian Language, 2001, Calcutta, India, 23 p. inria-00107878

**HAL Id: inria-00107878**

**<https://inria.hal.science/inria-00107878>**

Submitted on 19 Oct 2006

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

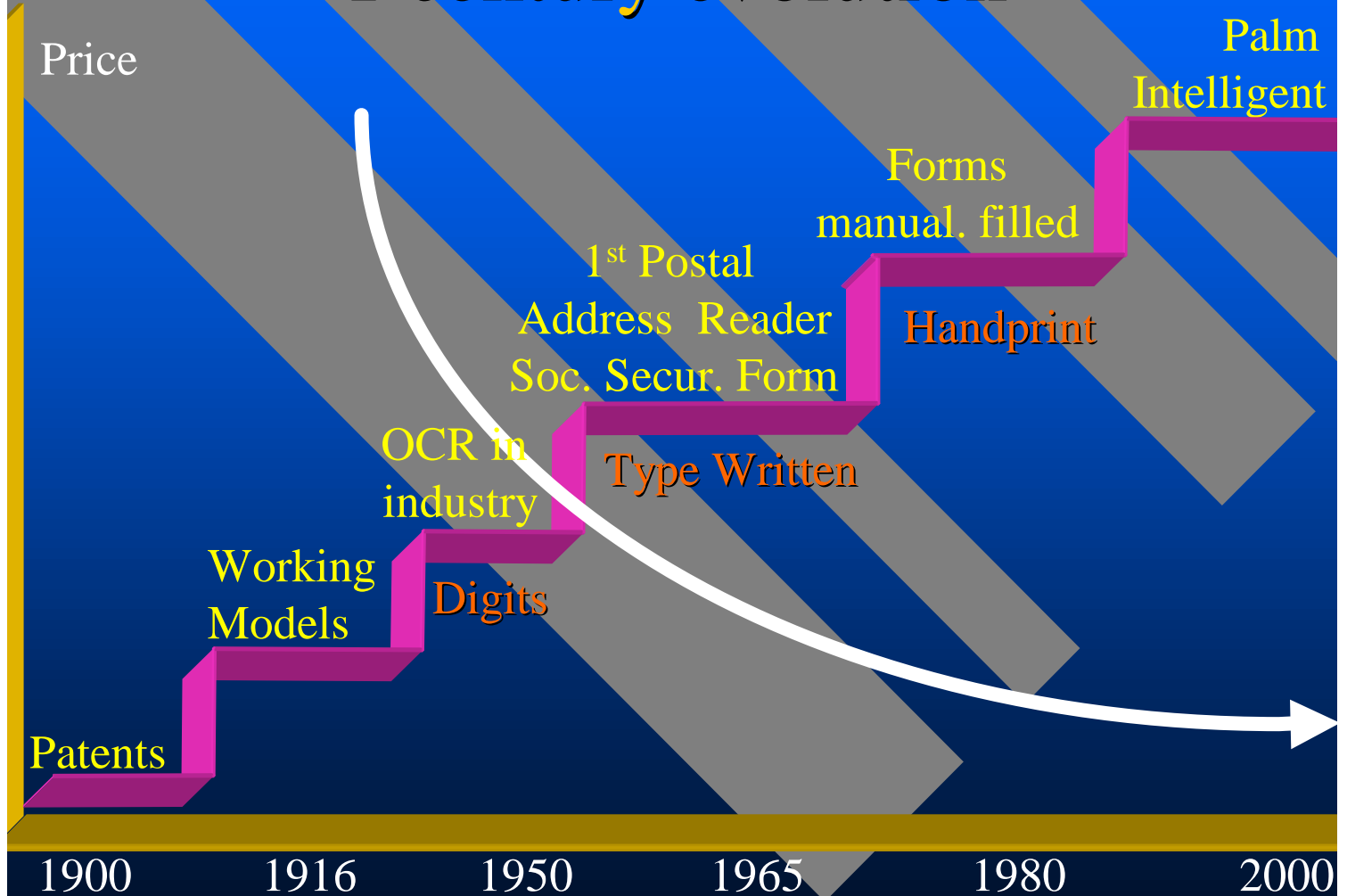
# A Short Tour of OCR, ICR, DIA

A. Belaïd

READ Group  
Nancy, France



# Historical 1 century evolution



# Handwritten-Printed

Handwriting

Approaches

- Training
- Recognition
- Contextual analysis

Printed

Applications

- Interface
- Postal au
- Bank che
- Forms

- Reading au
- Retro conv
- Translation

nnected  
storted (Free)

parated  
nt Style  
yout

# Handwriting:

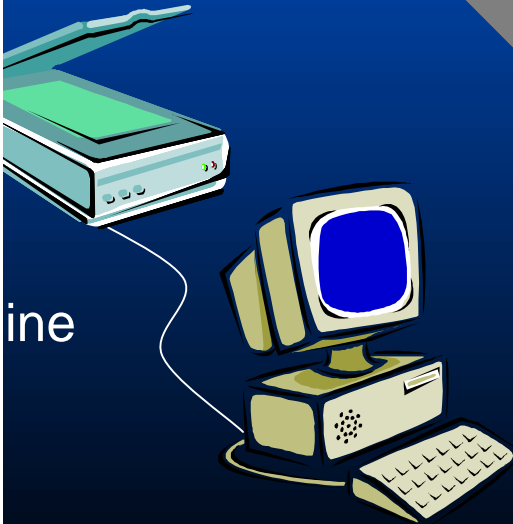
## Complexity : determinant factors

Acquisition

Spatial writing style [Tappert 1980]



On-line (Trackpad)



line

. Belaïd

BOXED DISCRETE CHAR

Spaced Discrete Characters

Run-on discretely written characters

pure cursive script writing

Mixed Cursive and Discrete

Writer Number

Mono, Multi, Omni-writer

Vocabulary Size

Very limited (<100 words), Extended (<10 000 words), Open

IWTDIL-2001, Calcutta 26/03/01

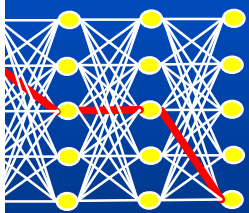
# Handwriting:

Methodology : different choices

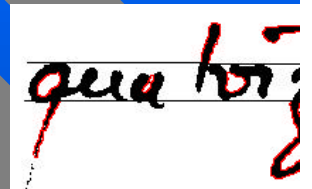
discriminant



discriminant



Global  
(holistic)

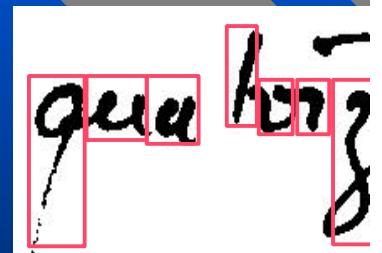


Training

Recognition

Segmentation

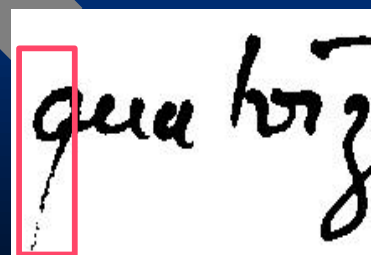
Analytical



External



Internal



Sliding Window

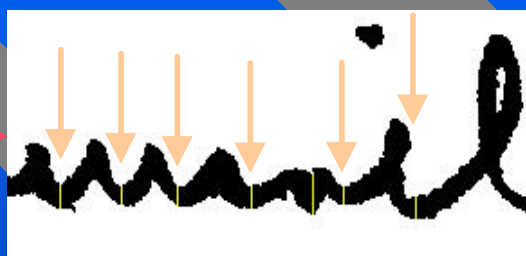
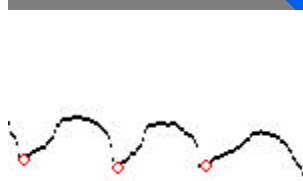
# Shape Analysis

As quick as thought is an old mode of expression used to convey an idea of the greatest rapidity; but no one until lately ever dreamed that a thought could be sent hundreds of miles in a few seconds, and that a person standing in London might hold a conversation with another in Edinburgh, put questions and receive answers, just as if they were seated together in one room, instead of being three hundred miles apart. The Electric Telegraph is another of the wonderful discoveries of modern times used to convey an idea of the greatest rapidity. But no one, until lately, ever dreamed that a thought could be sent hundreds of miles in a few seconds, and that a person standing in London might hold a conversation with another in Edinburgh, put questions and receive answers, just as if they were seated together in one room, instead of being three hundred miles apart. The Electric Telegraph is another of the wonderful discoveries of modern times.

# Segmentation

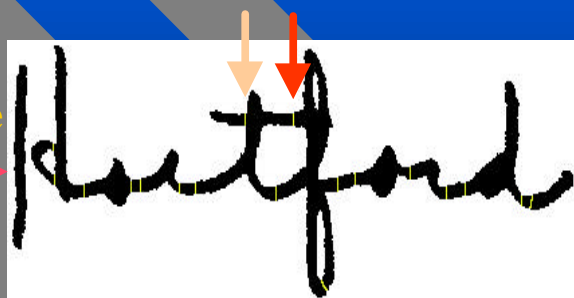
Profile Analysis [C. Scagliola, 2000]

Lower profile



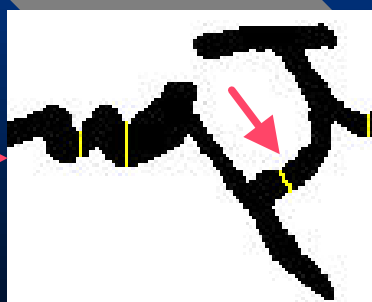
“Normal” ligature

Upper profile



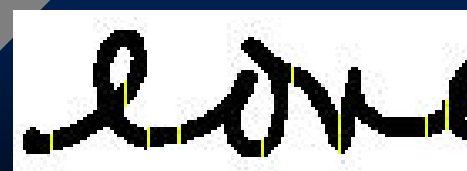
Letter contacts

Median profile



Hidden contacts

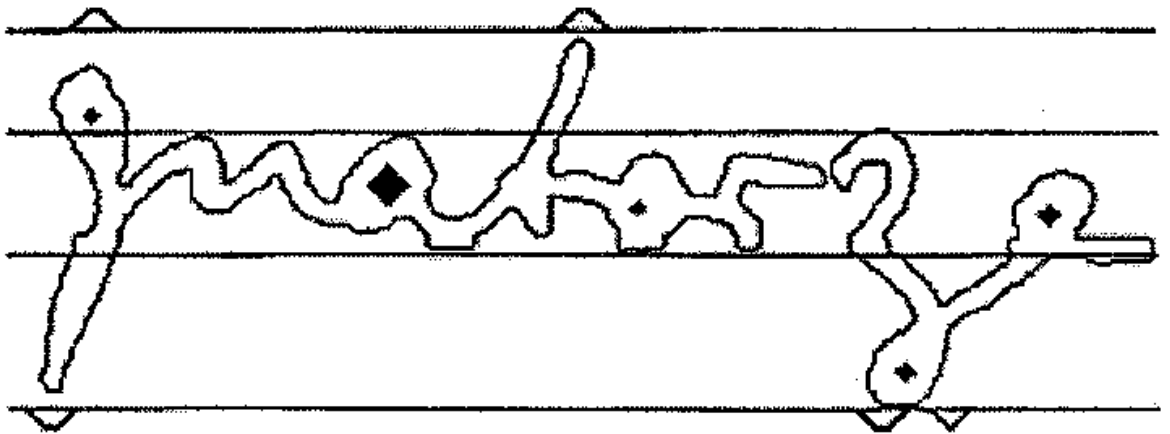
Single-run stretches : cut there are no minima and





# Recognition

quatorze



Global description : F — — OH — o — Zo  
by a string of symbols

# Recognition

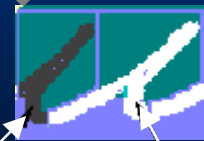
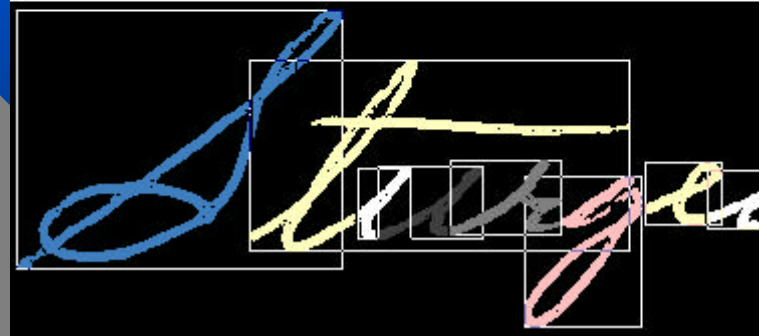
Hiterarchical [Lecolinet 1996]

Global

Analytical



representation of the image



1<sup>st</sup> grapheme

2<sup>nd</sup> grapheme



# Active Reading

## Xerox Xlibris project

- FX Laboratory Palo Alto

## High Resolution Tablet

- Display of the scanned image
- Writing on the screen with digital ink

## Reading Interface

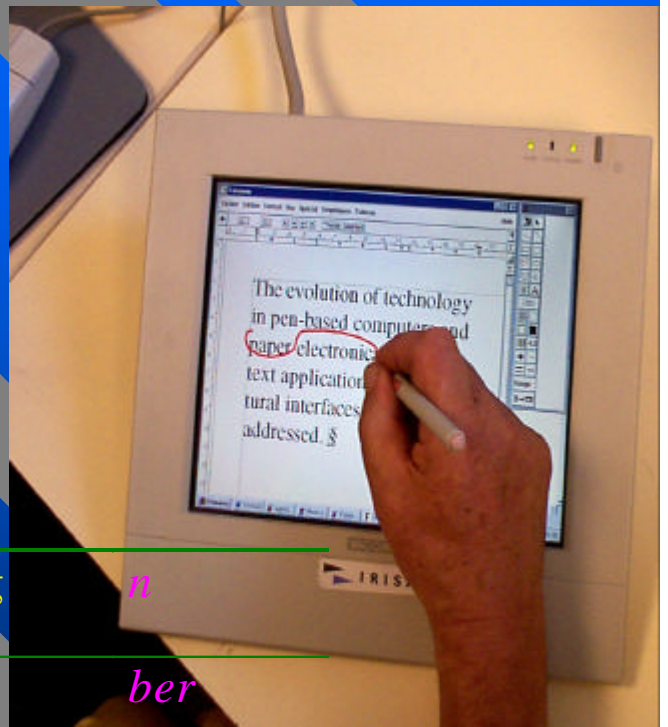
- Reading + action on the text :  
underlining, highlighting,  
commentaries in the margin  
or in a notebook



# Annotation and correction

## AmayaPen (IRISA) Graphical Interface

- Annotation
- Determination of the text zone



^ V	<del>Gute</del> berg	Gutenberg	n
e /	<del>Guten</del> ing	Gutenberg	ber
/	<del>Gutent</del> enberg	Gutenberg	
st	<u>Gutenberg</u>	Gutenberg	<i>Italique</i>
clavier, souris, etc.		souris, clavier, etc.	
e	<del>Gut</del> enberg	Gutenberg	

# Printed: Determinant factors

on Mode

Off-line / Scanner, Camera

Disposition

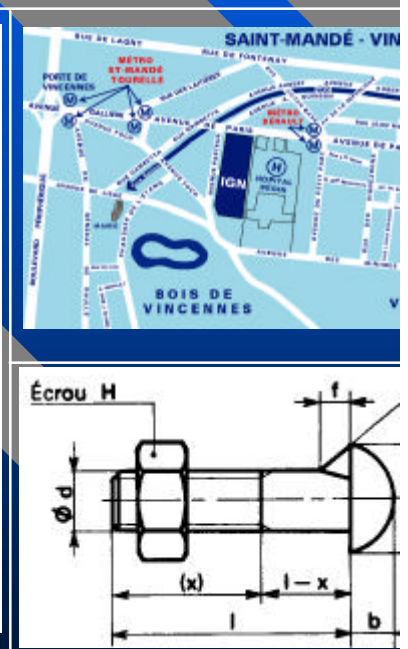
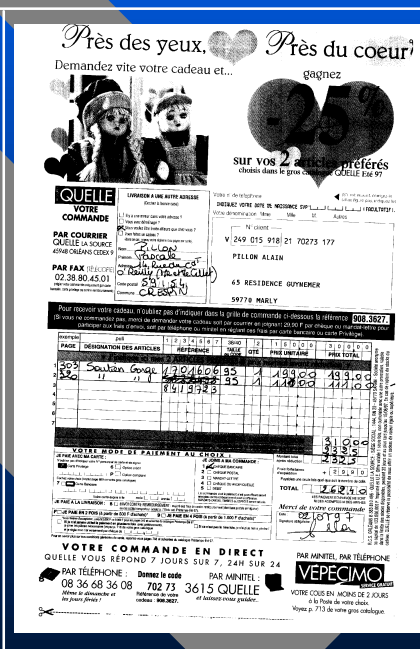
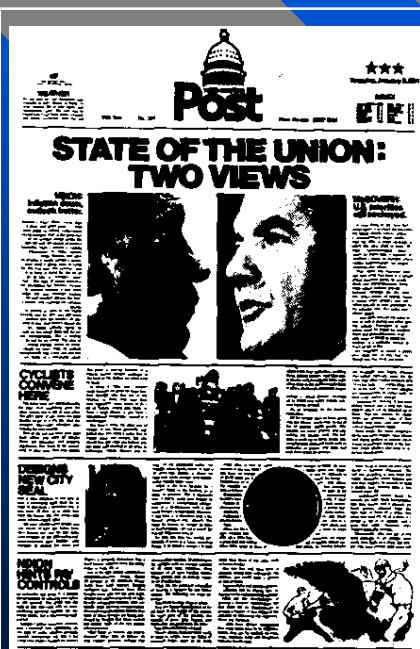
Variable according to the class

hattan

Mosaic

Zonal

Non linear



e Number

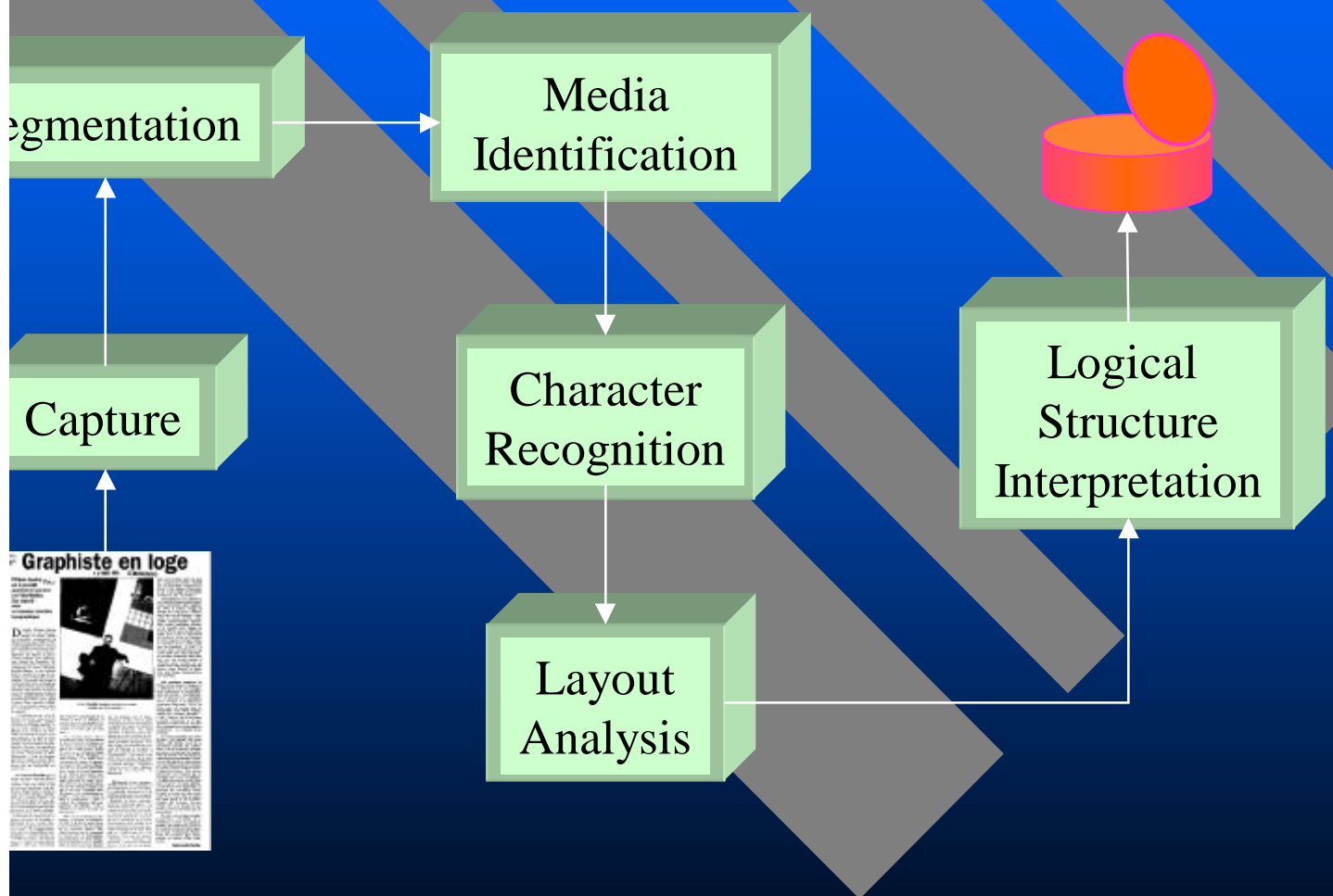
Mono, Multi or Omni-font

. Belaïd

IWTDIL-2001, Calcutta 26/03/01

# Recognition

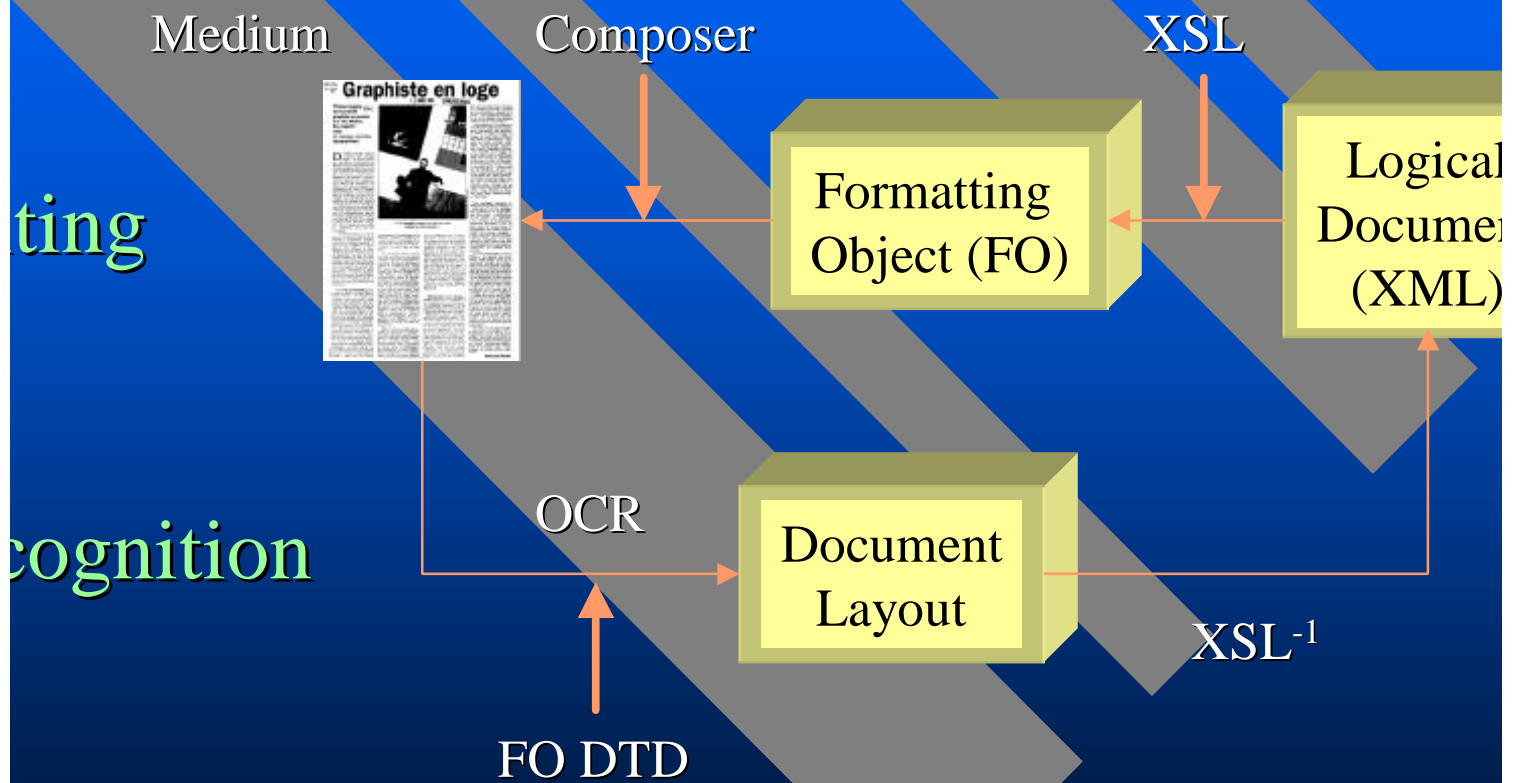
## The modules





# Recognition

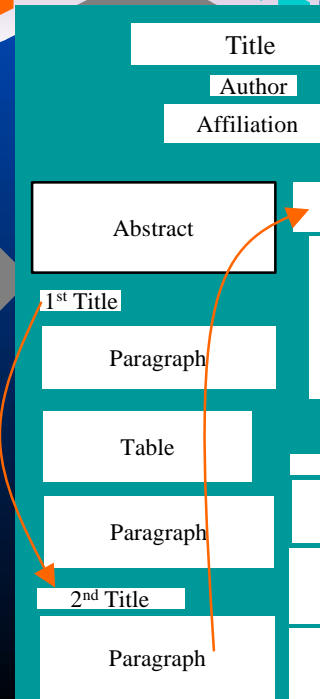
But recognition can be seen as the inverse editing process



Giving the idea to use some editing DTD for the recognition production process



## IWTDIL



# Automatic Generation

[Akindele-Belaïd 1994]

Specific Structure

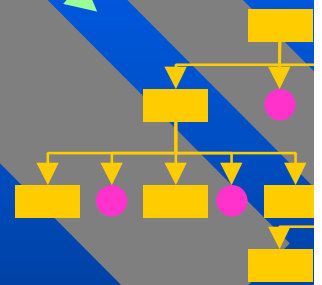
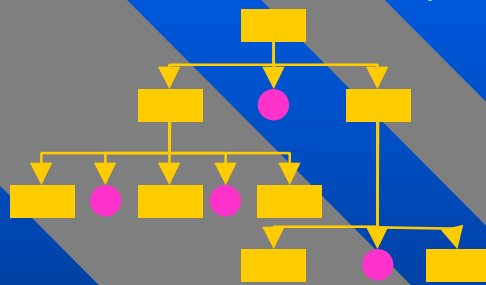
Logical Information

Basic Model

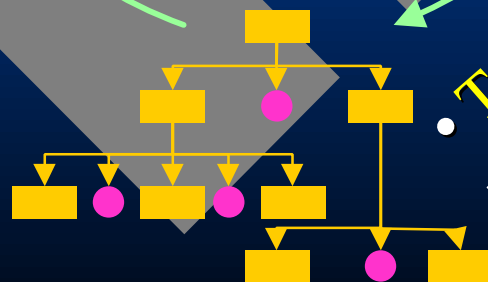
Construction of  
AGR, MOS

Layout

Physical-Layout



Generic Model




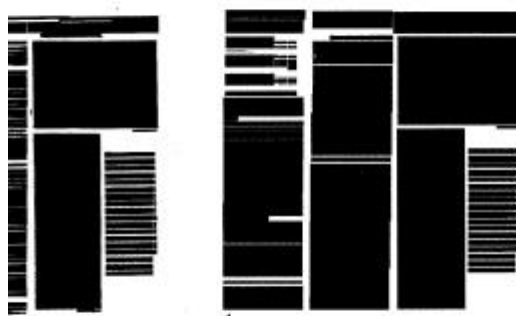
Tree Inference  
Constructor Transformation  
Attribute Update

Web

Online Proceedings

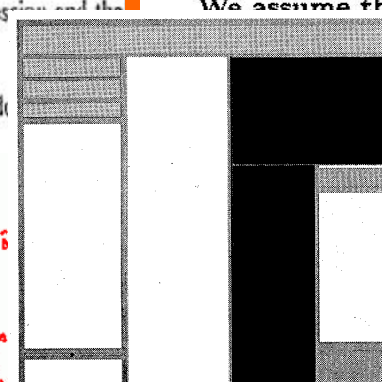
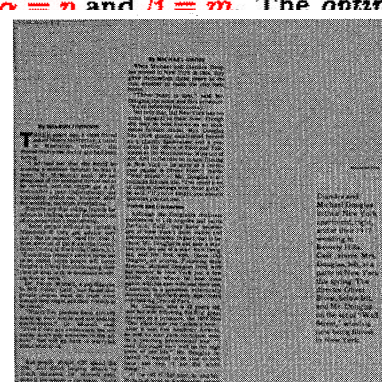
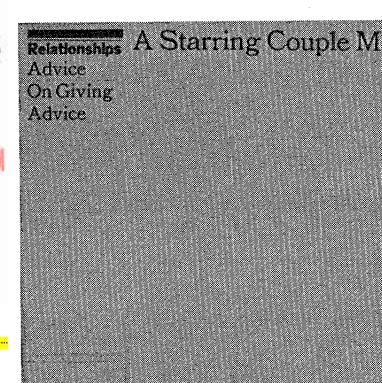
# Segmentation

Hough [Belaïd-Belaïd 1996] Fuzzy Sets [Kacem-Belaïd  
A [Wong 1982] Elements [Akindede-Belaïd 92]  
Fractals [Belaïd 1997] ~~Automatic Segmentation of Text~~


### 3. A BOOTSTRAP APPROACH FOR M SELECTION

We assume that  $\alpha = n$  and  $\beta = m$ . The optimal  $n$


TOTAL GENERAL (I + II + III + IV)

50 000

NJ

OK

$g_k(t)$

dis

$T -$

$X$

l u

$\sqrt{t}$

obs

$= 1,$

$=$

d in

$(t),$

$\sum_{k=1}^n \sum_{r=0}^m g_k(t) / r \cdot (t) \cdot (t)$

هذه الة

ة في الدر

موضوع

ن فقدان

المكتبة د

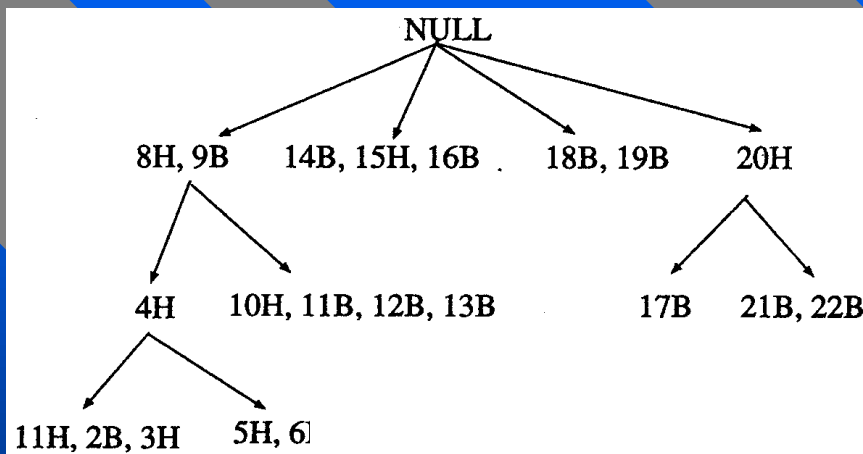
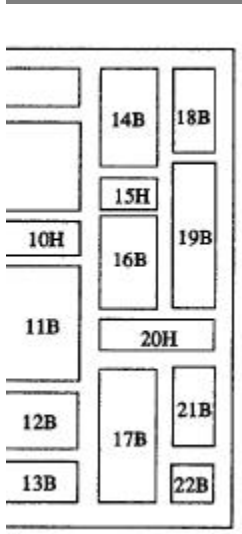
forme

استفعل

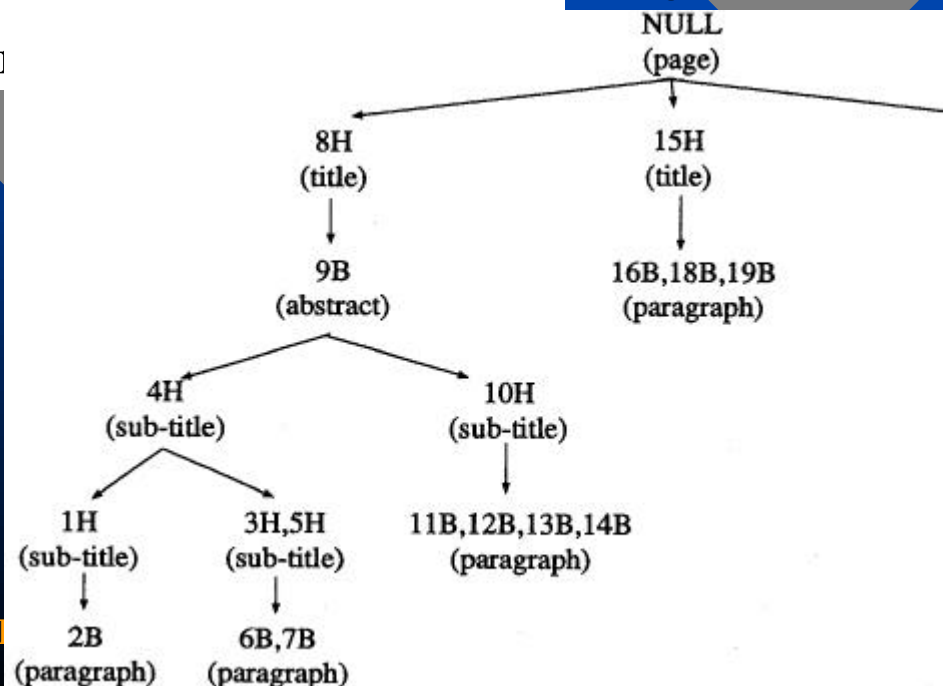
# Recognition

## Bottom-up Approach [Tsujiimoto 92]

### Geometrical Structure



### Logical Structure



and block  
by Block  
by ou Head

. Belaïd

IWT



# Recognition

## Top-Down Approach [Coasnon-Lorette 1998]


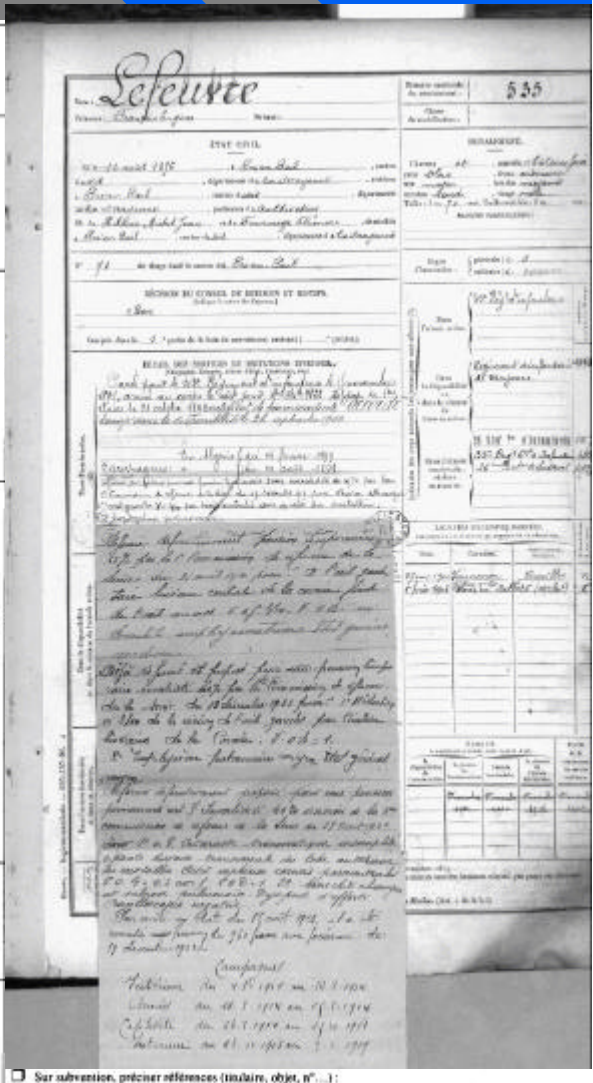
Positional grammatical Formalism

- » Document Description Language
- » Mixing syntactical & graphical description

Associated Analyzer

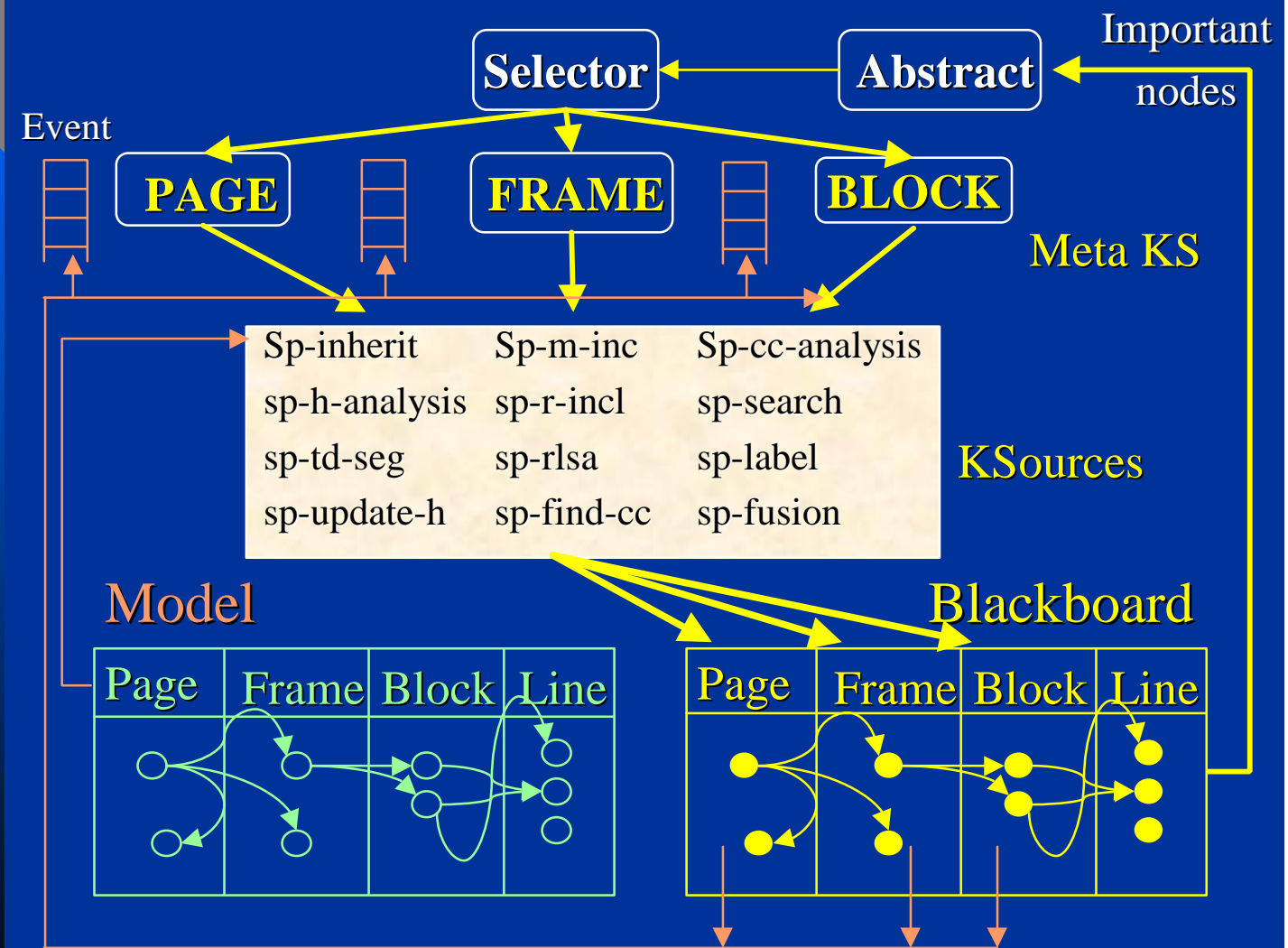
- » Reconsidering structure analyzed during analysis

$$\frac{2}{3} - \frac{\int_0^{\frac{x}{2}} y^2 + 3dy}{\frac{2}{3} + 56} - 1 + 2 \times$$

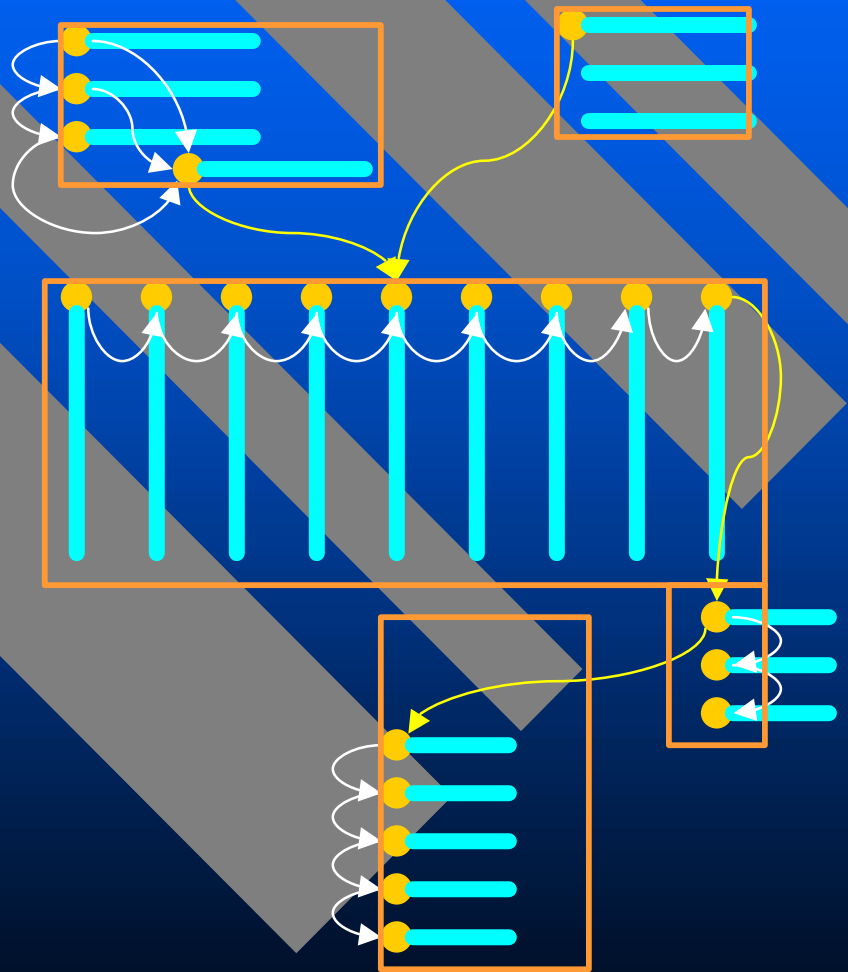
 INRIA Édition 1994	
<b>RÉFÉRENCES DU DEMANDEUR</b>	1. 10.10.1998 2. 10.10.1998 3. 10.10.1998 4. 10.10.1998 5. 10.10.1998 6. 10.10.1998 7. 10.10.1998 8. 10.10.1998 9. 10.10.1998 10. 10.10.1998
<b>OBJET DE LA MISSION</b>	1. 10.10.1998 2. 10.10.1998 3. 10.10.1998 4. 10.10.1998 5. 10.10.1998 6. 10.10.1998 7. 10.10.1998 8. 10.10.1998 9. 10.10.1998 10. 10.10.1998
<b>DÉROULEMENT DE LA MISSION</b>	1. 10.10.1998 2. 10.10.1998 3. 10.10.1998 4. 10.10.1998 5. 10.10.1998 6. 10.10.1998 7. 10.10.1998 8. 10.10.1998 9. 10.10.1998 10. 10.10.1998
<b>PRISE EN CHARGE DE LA MISSION</b>	1. 10.10.1998 2. 10.10.1998 3. 10.10.1998 4. 10.10.1998 5. 10.10.1998 6. 10.10.1998 7. 10.10.1998 8. 10.10.1998 9. 10.10.1998 10. 10.10.1998
<b>IMPUTATION</b>	1. 10.10.1998 2. 10.10.1998 3. 10.10.1998 4. 10.10.1998 5. 10.10.1998 6. 10.10.1998 7. 10.10.1998 8. 10.10.1998 9. 10.10.1998 10. 10.10.1998

# Recognition

## Multi-Agent Approach [Chenevoy-Belaïd 90]



[Valverde-Belaïd 1999]

[illegible]

# Conclusion

- **More and more paper !** In spite of the technology progress
- **New needs in the industry** With supplementary requirements
  - High Robustness ( —> 1 error / 10 000),
  - Fast Adaptation => changing structure
  - High Volumes => increase velocity, consider the overflow
- **Not only paper**
  - Document structure on the web
- **System to develop** should propose
  - Distributed architecture for DIA
    - » Useful structure (XML...), overflow solutions
  - Automatic production of meta-data
    - » making document readable and exchangeable on the Web